WestEd.org

Technology-Based Support Shows Promising Long-Term Impact on Math Learning:

Initial Results From a Randomized Controlled Trial in Middle Schools

Mingyu Feng
Chunwei Huang
Kelly Collins

WestEd

June 2023

Collaborators: WestEd, North Carolina State University, SRI International, LFC Research, Digital Promise, and Worcester Polytechnic Institute

Acknowledgments: We extend thanks to Linlin Li for her feedback on the research and Neil and Cristina Heffernan for their support during the study. We gratefully acknowledge the support of district leaders and the participation of the teachers and students within the districts where this study was conducted.

Suggested citation: Feng, M., Huang, C.-W., & Collins, K. (2023). *Technology-based support shows promising long-term impact on math learning: Initial results from a randomized controlled trial in middle schools.* WestEd.

This material is based on work supported by Arnold Ventures and the Institute of Education Sciences of the U.S. Department of Education under Grant R305A170641. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Arnold Ventures or the Institute of Education Sciences.

WestEd is a nonpartisan, nonprofit agency that conducts and applies research, develops evidence-based solutions, and provides services and resources in the realms of education, human development, and related fields, with the end goal of improving outcomes and ensuring equity for individuals from infancy through adulthood. For more information, visit WestEd.org. For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the E-Bulletin, our semimonthly e-newsletter, at WestEd.org/subscribe.

WestEd.org

Contents

Introduction 1

The Study 3

Study Design 3

Randomization 4

Implementation Context: Intervention Group 4

Implementation Context: Comparison Group 6

Data Collection and Analysis 6

Initial Results 8

Finding 1: ASSISTments Had a Significant Positive Impact on Student Learning in the Long Term 8

Finding 2: ASSISTments Helped Close Achievement Gaps 9

Finding 3: ASSISTments Benefited More Those Students Whose Schools Had a Higher Percentage of Students from Economically Disadvantaged Backgrounds or a Lower Percentage of White Students 11

Interpreting the Effects 12

Conclusions and Discussion 17

Ensuring Math Learning for All Students 17

Identifying the Successful Practices in Math Instruction and Learning 17

Building an Evidence Base for Technology-Based Programs 19

References 20

Introduction

Achievement gaps between demographic groups continue to be an important national and state-based concern (Bohrnstedt et al., 2015). The situation worsened during the COVID-19 pandemic. Mounting evidence documents how the pandemic massively disrupted K–12 students' education and increased educational inequity in the United States (Bacher-Hicks et al., 2020; Engzell et al., 2020; Maldonado & De Witte, 2020; Stein, 2020; U.S. Department of Education, 2021). The impact of the pandemic hit math the hardest, with some estimates indicating that the average student lost 5 months of learning (Dorn et al., 2021) and that achievement gaps widened significantly in math (Kuhfeld et al., 2020); learning loss is likely to be greatest among students from families of low incomes, Black, and Hispanic students (Center for Research on Education Outcomes, 2020; Dorn et al., 2020). In the most recent National Assessment of Educational Progress (NAEP) administration report card, math scores declined from prior years, reversing a long-term trend of score growth (U.S. Department of Education, 2022).

As a promising tool for improving math education and closing the achievement gap, the use of educational technology has dramatically expanded in K–12 education in recent years, accelerated by the COVID-19 pandemic. The current vision for educational technology is centered around creating personalized, engaging, and effective learning experiences for all students, regardless of their background. A wide variety of online resources and tools are available to support math education. These tools can provide students with instant feedback on their work, helping them understand their strengths and weaknesses in a particular subject or skill, correct their errors in real-time, and stay motivated in math learning. Teachers are provided with tools and resources, including digital assessments, learning management systems, data analytics, and professional development, to assess student understanding of math concepts, identify areas where students need additional support, and make data-driven instructional decisions accordingly.

Educational technology programs can improve student outcomes, but only if they are implemented effectively. Evidence-based practices can help ensure that technology is used in a way that will most likely benefit students. At the beginning of the COVID-19 pandemic, the U.S. Department of Education funded What Works Clearinghouse (WWC) to conduct a rapid review to synthesize existing evidence in online programs that promote learning. ASSISTments was one of the few digital learning programs recommended for use in response to the COVID-19 pandemic (Sahni et al., 2021). Developed by researchers at Worcester Polytechnic Institute (WPI), ASSISTments (Heffernan & Heffernan, 2014) is a free online formative assessment platform that applies innovative technology in education. The ASSISTments intervention aligns with theory- and empirically based instructional practices of formative assessment (Heritage & Popham, 2013) and skill development (Koedinger et al., 2013). It gives teachers new capabilities to assign math work and provides them with reports that summarize student work in real time to inform instruction adaptation. It also supports students' independent practice through immediate feedback and hints as they complete their assignments. With the impact of the COVID-19 pandemic, the use of digital platforms to support teaching and learning became necessary. As reported by the developer of ASSISTments, the use of ASSISTments in schools has increased significantly, from 800 to 20,000 teachers over the past 2 years (Heffernan & Heffernan, 2014).

The ASSISTments intervention aligns with theory- and empirically based instructional practices of formative assessment, data-driven instruction, and immediate feedback for students during independent practice.

ASSISTments' ability to help teachers improve math learning outcomes with their students has been, and continues to be, the subject of numerous studies. One particularly relevant study on the effectiveness of ASSISTments was conducted in Maine from 2013 to 2016 (Roschelle et al., 2016). The study sample consisted of 87 grade 7 math teachers and their students (totaling 2,850 students in the final study sample) from 46 middle schools. The study was a randomized controlled trial where randomization took place at the school level. Teachers remained in their schools' assigned condition for 2 years. Student outcomes were measured at the end of the second year using the TerraNova Common Core Mathematics assessment. An analysis was conducted using a hierarchical linear model (HLM) to analyze student outcomes by condition, and the results showed that students who used ASSISTments performed statistically significantly better than comparison students ($p < 0.001$, effect size Hedges' $g = 0.22$) (Murphy et al., 2020).

The Study

The follow-up study was designed to measure the possible long-term impact of ASSISTments on student learning, 1 year after the intervention was over.

From 2018 to 2020, WestEd led a randomized controlled trial in North Carolina to determine the replicability of the Maine study findings in a heterogeneous population that more closely matches the demographic diversity of the United States as a whole. The North Carolina replication study was designed to examine the efficacy of ASSISTments for advancing 7th grade students' math learning. Based on the study participants from the replication study, WestEd designed and conducted a follow-up study to measure the possible long-term impact of ASSISTments in 8th grade in 2021, 1 year after the intervention was over. The research team at WestEd investigated the extent to which the ASSISTments intervention supports students' math learning and whether the impact differs by subgroups. Two research questions motivated this follow-up study:

- RQ1: What is the long-term impact of ASSISTments on student math outcomes measured using the End-of-Grade (EOG) state test at the end of grade 8, 1 year after the completion of the intervention?
- RQ2: Do the long-term effects of ASSISTments, if they exist, vary for students of different socioeconomic status, race/ethnicity, or other policy-relevant characteristics?

This research report provides background of the North Carolina replication study and discusses the findings from the analyses of the follow-up study.

Study Design

For the North Carolina replication study, WestEd conducted a randomized experiment in 7th grade math classrooms across 63 schools. Schools were randomly assigned to either intervention or business-as-usual comparison conditions. ASSISTments was implemented by all grade 7 teachers in intervention schools over 2 consecutive school years—teachers "warmed up" and learned to use ASSISTments for a year in 2018/19, and then in 2019/20, they continued using ASSISTments with a new cohort of 7th grade students. Students who were in grade 7 in

2019/20 comprised the analytic sample for the follow-up study. In the follow-up study, to test the long-term impact of the ASSISTments platform in supporting math learning, students in the analytic sample maintained their original conditions and were followed for another year when their grade 8 performance (the long-term outcome) was measured in spring 2021 using North Carolina's 8th grade End-of-Grade Mathematics test. During the follow-up school year in 2020/21, no interventions were provided to grade 8 teachers or students by the developers or the research team. The follow-up study design and pre-analysis plan were pre-registered and published in Open Science Foundation Registries in 2019 (https://osf.io/exqpn).

The recruitment pool for the replication study consisted of schools in North Carolina, which was chosen due to its demographic diversity being very similar to that of the United States as a whole. Sixty-three schools from 41 districts took part in the study, including 48 Title I schools. Schools varied in the grades taught (i.e., K–8, 6–8, and 8–12) and the types of communities they served (i.e., rural, town, suburban, and city). A total of 102 grade 7 math teachers and their classrooms were enrolled in the study. A generalizability analysis based on Tipton (2014) suggested that the replication sample was highly representative of North Carolina schools and schools across the nation (generalizability index = 0.899, indicating high generalizability).

Randomization

Schools within each district were paired based on demographic characteristics, including their 7th grade enrollment, percentage of students from economically disadvantaged backgrounds, percentage of Black and Hispanic students, and student prior performance on state math assessments in the previous 2 years. Schools were randomly assigned to the intervention or comparison group within each pair. Randomizing within pairs ensured that both the intervention and comparison groups included schools from all participating districts. All grade 7 teachers in a given school had the same assignment—they either implemented the ASSISTments materials (intervention group) or the business-as-usual materials (comparison group) during the 2018/19 and 2019/20 school years.

Implementation Context: Intervention Group

During the study, the intervention was positioned as online support for math homework, a key opportunity for student independent practice, and for teachers to use formative data to adapt instruction. To prepare teachers in the intervention condition to effectively implement ASSISTments with their students, the ASSISTments development team provided systematic training and support to teachers. The ASSISTments professional development consisted of a 2-day in-person session before the start of each school year (both summer 2018 and 2019). Along with helping teachers learn and understand the technical features of ASSISTments, training also focused on how teachers could use information from reports to inform their instructional decisions, such as identifying which concepts students might need extra time to learn or which concepts students grasped more quickly. Teachers were also trained to use

ASSISTments to encourage students to rework problems they initially got wrong and to identify and discuss commonly missed questions or common wrong answers with students to address underlying misunderstandings. To support the ongoing implementation of ASSISTments, a local coach, who had experience as a former high school teacher and organizer of math and English remediation programs, visited participating teachers' classrooms in person three times each school year and provided teachers with continued guidance and assistance in the use of ASSISTments and how to integrate it into their classroom instruction.

> *Teachers were trained to use ASSISTments to encourage students to rework problems they initially got wrong and to identify and discuss commonly missed questions or common wrong answers with students.*

ASSISTments emphasizes a four-step loop that can be readily integrated into day-to-day instructional practices. The intensive professional development and continuous support also reinforce teachers' implementation of the practices. The four steps are

1. teachers find and assign content in the system that is aligned to the focus of class instruction and the state standards;
2. students work on assignments in the system and receive instant feedback and support;
3. teachers examine student performance data to inform instructional decisions; and
4. teachers share and review data with students in class and engage in discussion of common misconceptions.

By following these steps, ASSISTments enables teachers to make data-driven instructional decisions, facilitates formative assessment, and allows students to have supported independent practice.

The ASSISTments platform contains textbook-based problems from the school's existing textbooks and curricula, as well as other prebuilt content available for teachers to assign to students through the ASSISTments platform. The ASSISTments development team also built practice problems created by the intervention teachers into the platform upon their request. During the study, intervention teachers were expected to assign 20–30 minutes of ASSISTments assignments at least twice a week and were guided to regularly review data reports of student work on assignments.

Implementation Context: Comparison Group

The comparison group teachers were asked to use their regular curriculum materials and continue with their existing homework practices as they usually would. They were also asked to maintain their regular practices in regard to attending any professional development workshops or receiving any math coaching typically provided by their district. All teachers in both the intervention and comparison groups in the study were expected to follow the North Carolina Standard Course of Study (NCSCOS). At the time of recruitment and randomization, North Carolina schools were at various stages of technology adoption. Many teachers were already using online platforms such as IXL, iReady, and Khan Academy to assign homework to their students. Comparison teachers continued with their existing use of technology tools during the course of the study but had no access to ASSISTments.

After the conclusion of the 2019/20 school year, grade 7 teachers in the comparison group were invited to participate in an optional 2-day professional development session and were provided access to ASSISTments to use in their classrooms starting in fall 2020. A small group of teachers joined the professional development session, but no implementation support was provided by the development team or research team during the 2020/21 school year to either intervention or comparison groups.

A Note About the Broader Study Context

The North Carolina replication study was in its second year of implementation during the 2019/20 school year when the COVID-19 pandemic caused school closures in March 2020. The North Carolina End-Of-Grade (EOG) state standardized test was subsequently cancelled for spring 2020. ASSISTments usage was interrupted due to school closures and then resumed in approximately half of the intervention teachers' classrooms during remote learning. The grade 7 state EOG test scores were missing for all students in the study during the 2019/20 school year.

Data Collection and Analysis

The data analysis followed the pre-analysis plan. We obtained demographic, enrollment, and state assessment performance data from the statewide database through the North Carolina Education Research Data Center (NCERDC) for all students in the 63 study schools. The follow-up study used the North Carolina state standardized End-of-Grade (EOG) math assessment for grade 8 (MA08) to measure the students' long-term learning outcomes. Students' immediate learning outcome in grade 7 was missing because the statewide testing in spring 2020 was canceled due to the COVID-19 pandemic. Grade 6 Mathematics test (EOG-MA06) scores from spring 2019 served as the baseline measure. The North Carolina EOG math assessments were designed based on the Common Core State Standards for Mathematics (CCSSM) to measure students' proficiency since 2010. These assessments are administered annually to students in grades 3 through 8.

Individual demographic data included gender, race/ethnicity, and economically disadvantaged status (EDS). School-level background data included average EOG-MA06 score, 7th grade enrollment size, Title 1 eligibility, percentage of students with EDS, and percentage of ethnic groups. These variables served primarily as covariates in the impact analyses.

We tested the baseline performance equivalency between intervention and comparison groups using their MA06 test scores. The difference is small and not statistically significant at the 0.05 level (545.70 for intervention versus 545.03 for comparison, Hedges' g = 0.08), indicating that both groups performed equally at the end of 6th grade prior to participating in the study.

To conduct the impact analyses, an intent-to-treat (ITT)[1] approach was adopted using a student sample from all 63 original schools. Students were included in the study sample if they had both 6th grade (from the 2018/19 school year) and 8th grade (from the 2020/21 school year) North Carolina End-of-Grade (EOG) scores and if they were enrolled in one of the 63 participating schools as a 6th grade student during the 2018/19 school year or as a 7th grade student during the 2019/20 school year if the school did not serve 6th grade. The student sample included in this report consisted of 5,991 students who took EOG-MA08, including 2,961 students in the intervention group and 3,030 students in the comparison group. All 63 schools were represented in the sample. Therefore the school level attrition was 0 percent. Another group of 1,962 students took accelerated or alternative math courses and did not take the EOG-MA08 test. Their scores were analyzed separately, and the findings were not included in this research report.

[1] An intent-to-treat (ITT) analysis is one in which all participants originally randomized at the onset of the study are included in the analysis in their original randomization groups. They are included regardless of whether or not they were able to maintain participation in their original randomization groups throughout the entire study.

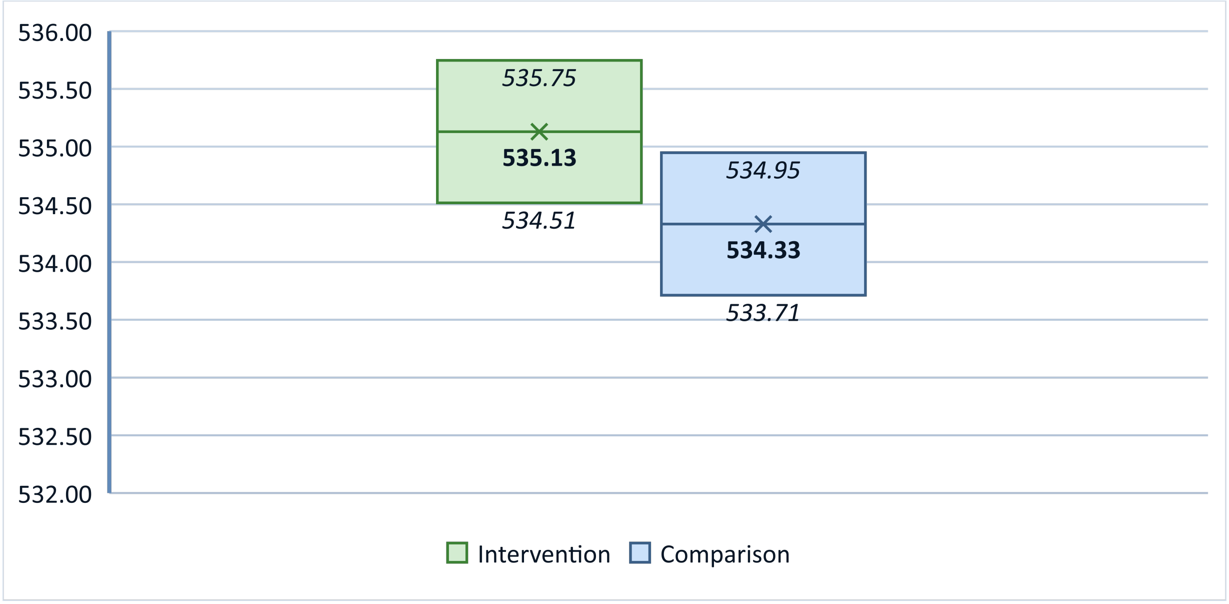
Initial Results

Three main findings from the follow-up study show that ASSISTments (1) had a positive impact on student learning in the long term, (2) helped close achievement gaps between White students and students of color, and (3) benefitted more those students whose schools had a higher percentage of students from economically disadvantaged backgrounds or a lower percentage of White students.

Finding 1: ASSISTments Had a Significant Positive Impact on Student Learning in the Long Term

To estimate the long-term impact of ASSISTments on student learning outcomes, we compared the EOG-MA08 scores for students in intervention schools with those in comparison schools. The analysis used two-level hierarchical linear regression models (students nested within schools) controlling for school-level and student-level characteristics and students' prior achievement on their EOG-MA06 tests. The initial analysis showed that students in the intervention schools assigned to use ASSISTments in 7th grade scored 0.8 points higher than comparison students on the EOG-MA08 test. The estimated mean score on MA08 for the intervention group was 535.13 (standard deviation = 8.49) versus 534.33 for the comparison group (standard deviation = 7.72) (Figure 1). This treatment impact is statistically significant ($p = 0.011$) and corresponds to an effect size of 0.10 (Hedges' g). The results suggest that implementing ASSISTments had a significant positive impact on student learning in math in the long term. The effect was sustained after 1 year of program implementation, which ended at the end of 7th grade.

Figure 1. Estimated Mean Score on End-of Grade Math Assessment for 8th Grade: Intervention Versus Comparison



Note. **The chart shows the students' MA08 estimated mean score (bolded numbers) with standard error (the italicized numbers which represent upper and lower bounds of the standard error) for both the intervention and comparison groups, respectively.**

Finding 2: ASSISTments Helped Close Achievement Gaps

We also conducted a series of analyses to study the differential impacts by subgroup. The analyses based on student-level backgrounds indicated that

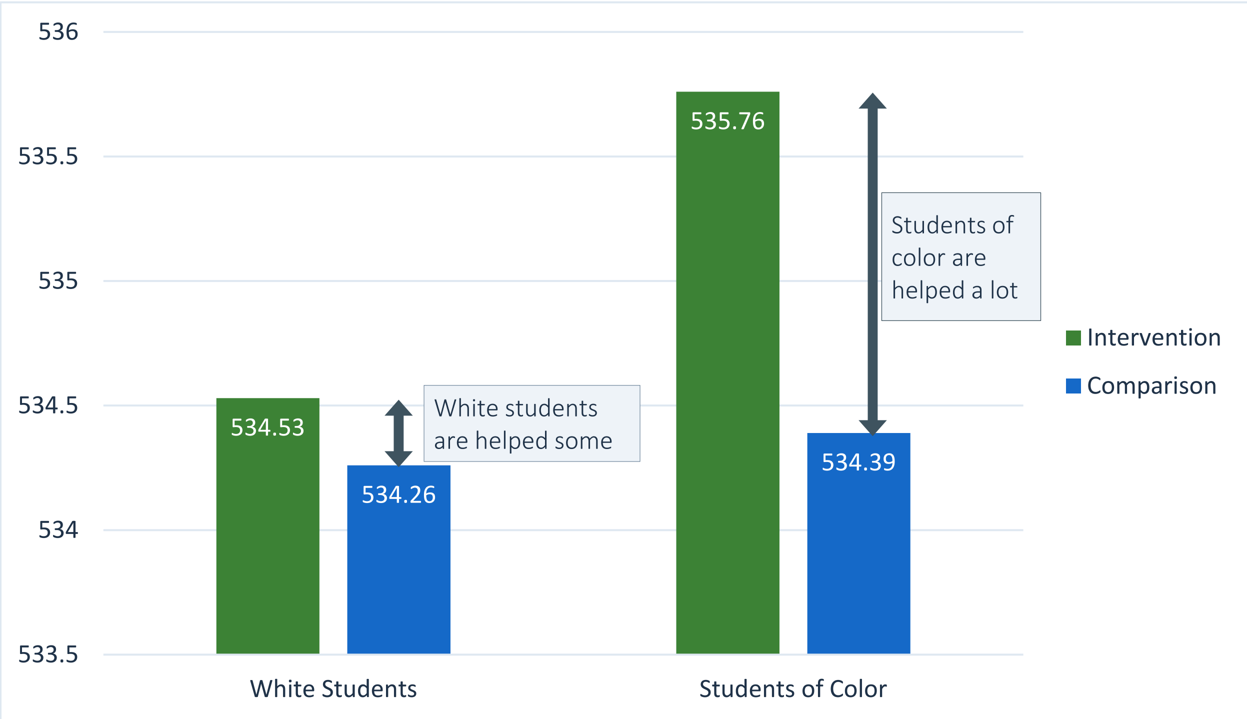
- the intervention benefited students of color significantly more than White students (p = 0.003, Hedges' g = 0.14), and
- implementing ASSISTments benefited Hispanic students more than non-Hispanic students (p = 0.014, Hedges' g = 0.13).

At the baseline in either the intervention or comparison group, the students of color scored 2 points lower than White students on MA06 (Table 1). However, at the end of 8th grade, the students of color scored higher than White students in the intervention group versus no difference in the comparison group (Figure 2). Both students of color and White students in the intervention group scored higher than their counterparts in the comparison group. The results suggest that traditionally underrepresented groups (such as Hispanic students) benefited more from the ASSISTments program than did White students. Therefore, the intervention helped narrow the achievement gap among students of different ethnicities.

Table 1. ***Baseline Difference of Average Score on MA06: White Students and Students of Color***

	Intervention		Comparison	
	mean score	***n***	**mean score**	***n***
White students	546.56	1,501	546.27	1,556
Students of color	544.72	1,465	543.79	1,474
Difference	1.84		2.48	

Figure 2. ***Average MA08 Score by Experimental Condition and Ethnicity (White versus Students of Color)***



***Note*. Both White students and students of color in the intervention group scored higher than their counterparts in the comparison group. However, at the end of 8th grade, the students of color scored higher than White students in the intervention group versus no difference in the comparison group.**

Finding 3: ASSISTments Benefited More Those Students Whose Schools Had a Higher Percentage of Students from Economically Disadvantaged Backgrounds or a Lower Percentage of White Students

We also examined the differential impacts of school-level backgrounds to see whether the effects of ASSISTments vary based on school settings or its composition of student populations. We split the schools into two groups based on the average percentage of students from economically disadvantaged backgrounds and the average percentage of White students. The results suggested that students in intervention schools with a higher percentage of economically disadvantaged students benefited more than those in intervention schools that had a lower percentage of students from economically disadvantaged backgrounds ($p = 0.045$, Hedges' $g = 0.22$). Furthermore, students in the intervention schools with a lower percentage of White students seemed to benefit more than those in the intervention schools with a higher percentage of White students ($p = 0.056$, Hedges' $g = 0.19$). Studies have shown that school districts that serve more Black and Latino students and students from families of low incomes were on average receiving less funding than school districts that serve more White students and higher-income counterparts (Morgan, 2022). The results demonstrated the potential positive impacts of the ASSISTments program in providing needed support to students and teachers in schools where access to educational resources might be limited.

Interpreting the Effects

While the findings described above are significant, it is important to view them in the real-world context of what the differences in scores mean for improvement in math learning of various groups of students.

An effect size is a common way in education research to quantify the magnitude of the difference between two groups. To put this in perspective, we considered several approaches to help contextualize the impact on student learning and explain the effect size, building on the guidance of Lipsey et al. (2012), a leading researcher who developed broad recommendations for effect size reporting and interpreting.

First, we referred to the achievement level ranges for North Carolina’s statewide End-of-Grade tests of Mathematics (North Carolina Department of Public Instruction [NCDPI], 2022b). Table 2 displays the achievement level cut scores for MA08. The table suggests that the score range within one achievement level tends to be narrow and within just a few points. For instance, the score range for Level 3 is only 5 points, and the score range is 7 points for Level 4, which suggests that an increase of 0.8 points in the intervention group could be a substantial improvement.

Table 2. ***North Carolina Grade 8 Math End-of-Grade Test Achievement Level Score Ranges Versus Average Mean Score for the Experimental Groups***

	Not Proficient	Level 3	Level 4	Level 5
General education mathematics	542 and below	543–547	548–554	555 and up
Mean score for intervention group	535.13			
Mean score for comparison group	534.33			

Note. **This table shows the score cutoffs for various categories of proficiency on the EOG assessment for math Grade 8. Level 3 indicates that a student meets on-grade-level proficiency, and Levels 4 and 5 indicate that a student meets college and career readiness standards. While the impact is significant, we observed that the estimated mean score for the intervention group is still 8 points short of being "Proficient" on grade 8 standards (535 versus 543, a cut score for being categorized as Proficient at Level 3), which highlights the importance and urgency of supporting students to recover from the learning loss induced by the pandemic.**

Second, Lipsey et al. (2012) suggested that it is helpful to compare the size of the effect to a well-known benchmark that is relevant to policymakers. The persistent achievement gap between students of color and White students is one such benchmark that policymakers and administrators are aware of (Schneider, 2022). Using that benchmark, one could ask, "If the overall effect size of the intervention is 0.1, what percent of the achievement gap would it help resolve?" To answer this question, we analyzed the 2020/21 EOG-MA08 test results from all 189 schools that have 8th grade in North Carolina. Table 3 shows the statewide average scores overall and the mean scores by ethnic group. We found that the statewide mean score in 2020/21 was 533.95, very close to the estimated mean score of 534.33 for the comparison group in the study. We noted that the mean score for Black students was 2.79 points below the statewide average, and the estimated difference between the intervention group and the comparison group was almost 30 percent of the discrepancy (0.8 versus 2.79).

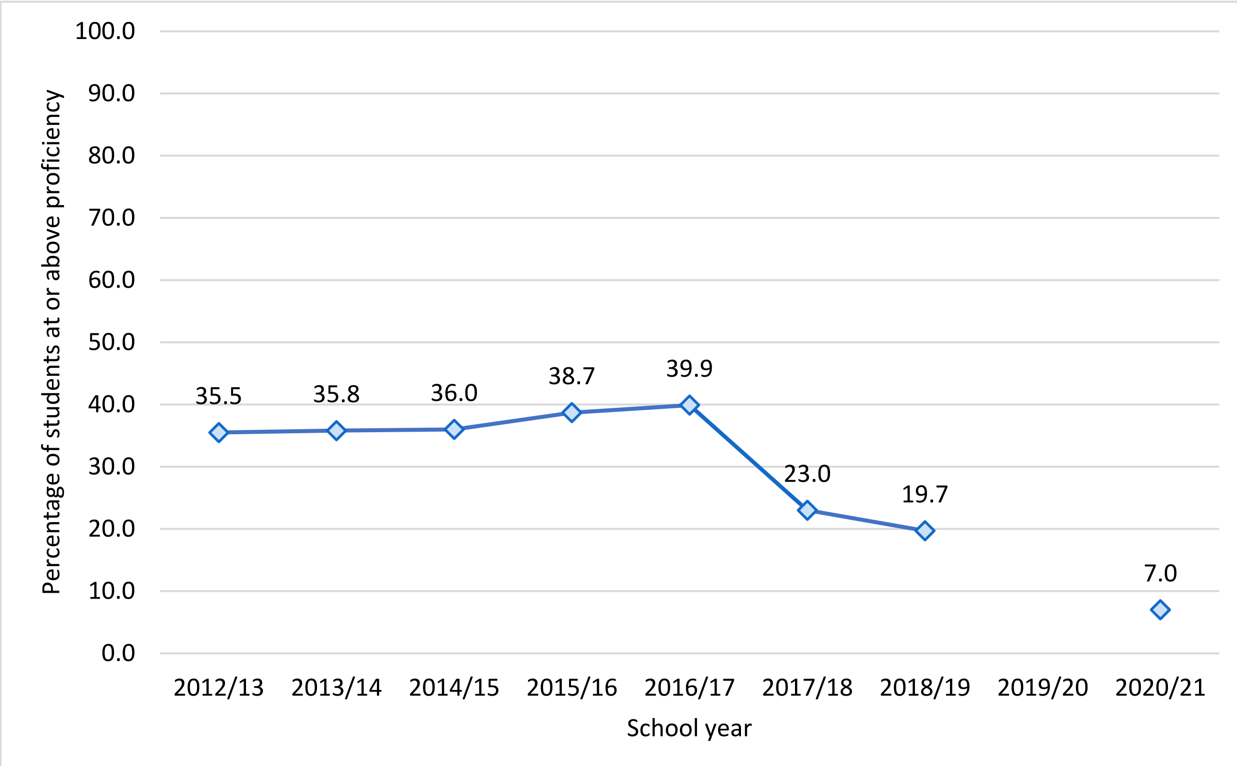
Table 3. ***North Carolina Statewide EOG-MA08 Score Descriptive Statistics by Ethnicity Groups***

Group	*n* (Students)	Mean score	Standard deviation
Statewide overall	79,093	533.95	7.95
Black students only	22,162	531.16	6.48
Hispanic students only	17,876	533.19	7.54
White students only	32,584	536.17	8.31

Third, we examined the historical trends in the North Carolina EOG assessment and compared the results from 2020/21 with results from prior years to help us interpret the difference between the two experimental groups (NCDPI, 2022a). We observed that the statewide mean test scores decreased by 4.6 points from 2018/19 (538.6 with 19.7 percent of students categorized as Proficient) to 2020/21 (534 with 7 percent of students categorized as Proficient), which is likely due to the learning loss caused by the COVID-19 pandemic.

Figure 3, with data from the NCDPI Historical Trends and Results Report for the 2020/21 school year (2022a), illustrates the dramatic drop in the percentage of students who tested at or above Level 4 in math post-pandemic between the 2018/19 EOG tests and the 2020/21 EOG tests. Unfortunately, because of the pandemic, no data for 2019/20 is available. We therefore could not compare the results from the school years before 2018/19 (the new performance levels were set and became effective during the 2018/19 school year). Despite the limited information available to inform the long-term trend of the North Carolina EOG test, the improvement found in the intervention group might still be practically meaningful; students receiving ASSISTments intervention experienced 17 percent less of the learning loss (0.8 out of 4.6 points) than those in the comparison group or other North Carolina schools statewide on average.

Figure 3. ***Percentage of North Carolina 8th Grade Students Testing at or Above Proficiency in Math Over Time***



Note. **The first year Common Core Standards were implemented was the 2012/13 school year.**
For the 2012/13 school year, "At or Above Proficiency" was calculated by dividing the number of students passing math at or above Level 3 by the number of students with valid scores.
For the 2013/14 school year and later, "At or Above Proficiency" was calculated by dividing the number of students passing math at or above Level 4 by the number of students with valid scores.
New standards for math were implemented during the 2018/19 school year (NCDPI, 2022a).

Although there is limited information to examine the long-term trend of the North Carolina End-of-Grade test, the 0.8-point increase found in the intervention group could still be practically meaningful.

Fourth, we compared the results with similar studies or interventions as reported in meta-analyses. We found that the long-term effect size of the current study is about the same as the immediate impact effect sizes found from those studies or interventions in the meta-analyses. The overall effect size of 0.1 in the current study is about the same as the reported mean effect size for the "curriculum or broad instructional program" type of intervention (Lipsey et al., 2012, Table 10, p. 36). Lipsey et al. included only 13 middle school randomized study (immediate) effect sizes that used broad-scope standardized tests as student learning outcome measures, and the median effect size was 0.11 (mean = 0.15, Table 9). Similarly, Kraft (2019) reported a mean effect size of 0.11 (or 0.07 for median effect size) from randomized control trials of education interventions with standardized achievement outcomes for grade 8 math (Appendix B, p. 42). This indicates that long-term effects of the intervention on student learning 1 year after the ASSISTments intervention ended are comparable to the immediate effects of other similar interventions.

Regardless of the grade level or subject, the mean effect size was only 0.05 for the 124 studies with a sample size larger than 2,000 (Table 1, p. 34), and 80 percent of the effect sizes reported in these large-scale studies were below 0.11.[2] In addition, Kraft (2019) found that the mean (or median) effect size was 0.03 from those 49 studies funded by the U.S. Department of Education. These effect size numbers suggest that the effect size of 0.1 reported in the current study should not be considered small, given that it is a large-scale study, based on the broad achievement outcome measured by the state annual testing program, and the 0.1 effect size represented a long-term effect 1 school year after the intervention was completed.

It should be noted that few studies have reported a long-term effect of an intervention such as the current study. It is generally expected that the effects of interventions fade away as time passes, and the long-term effect would be smaller than the short-term effect measured right after the end of the first-year intervention. Adding this point of view, the effect size of the current study for the long-term impact is particularly notable because it indicates that the immediate effect was greater than those from other comparable interventions.

[2] Many of the interventions that demonstrated significant effect sizes in the large-scale studies were high dosage tutoring programs that are costly. Guryan et al. (2023) reported the tutoring cost by Saga was $3,200 to $4,800 per year per student for a comparable delayed effect of 0.10 standard deviation on a standardized test. We are still in the process of conducting the cost analysis for the ASSISTments implementation during the North Carolina replication study. Initial rough estimate of the cost of ASSISTments per student per year was under $100.

Conclusions and Discussion

Technology-based programs, along with accompanying professional learning and integrated with practices, can support teachers in creating classroom conditions that better prepare students for future math learning.

Math education is a crucial aspect of academic development, and closing achievement gaps is a pressing issue in education today. This study and findings contribute to an increased understanding of the potential of technology-based solutions for advancing math learning and recovering from learning loss caused by the COVID-19 pandemic.

Ensuring Math Learning for All Students

The study findings have significant implications for improving math learning in middle school for *all* students through technology-based supplemental programs aligned to class instruction. The research was conducted in schools with diverse student populations, closely reflecting the makeup of student populations across the nation. Notably, the study provided empirical evidence that students whose teachers used ASSISTments for homework support performed significantly better in the state-administered general assessment 1 year later. The impact of the intervention was particularly pronounced for students from historically underrepresented racial and ethnic groups in STEM education and careers (Fry et al., 2021) and those attending traditionally under-resourced schools. This highlights the potential of technology-based programs to reduce achievement gaps. These findings underscore the importance of scaling up effective technology-based programs to promote equitable access to high-quality math education for all students and prepare them for future math learning.

Identifying the Successful Practices in Math Instruction and Learning

The results from this study show that technology-based programs, along with accompanying professional learning and integrated with practices, can support teachers in creating classroom conditions that better prepare students for future math learning. Despite being influenced by COVID-19, this study is among the limited number of experimental research studies on math educational technologies in which

- the research is a large-scale replication of a prior study;
- the study examined a universal intervention that was implemented schoolwide with a broad student population, as opposed to programs that target a selected subset of students;
- the student learning outcomes were measured by an end-of-year standardized math assessment included in the state accountability program, as opposed to a narrow assessment that focuses on unit skills or knowledge; and
- students were followed across grades to examine long-term retention effects.

Roschelle and Hodkowski (2020) suggested using research to inform practices during the pandemic and noted the importance of feedback and adaptiveness. We hope the findings from this study will inspire educational researchers to conduct more targeted studies to further investigate the specific mechanisms and instructional strategies within ASSISTments that led to the positive outcomes.

It is important to recognize that teachers in the comparison group used a variety of supplemental technology programs intended to support students' progress toward grade-level learning standards. This highlights the need to identify the specific ingredients of ASSISTments that contributed to the observed differences in learning outcomes between the intervention and comparison groups, which could include formative assessment (i.e., using data to drive instruction decisions), opportunities for student independent practice with immediate feedback and assistance, prebuilt standard-aligned content, and so on.

It is important to recognize that teachers in the comparison group used a variety of supplemental technology programs intended to support students' progress toward grade-level learning standards. ASSISTments, and its accompanying professional development, need to be further studied to identify the specific ingredients of the program that contribute to observed differences in student learning outcomes.

Building an Evidence Base for Technology-Based Programs

The study highlights the essential role of research in building an evidence base for educational technologies and calls for attention to evidence-based programs and practices.

Numerous technology-based learning platforms and programs have been built for K–12 classrooms. As these materials become more widely used across different geographic regions and with varying student populations, it is critical to investigate their impact on student learning in the short and long term, taking into account student diversity, equity, and school context. Conducting rigorous research in authentic school settings helps ensure that educational technologies are effectively implemented to meet the needs of all students and advance equitable educational outcomes. This is especially true since educational technology programs have the potential to widen the achievement gap if they are not implemented in an equitable manner. Unequal access, lack of data literacy and teacher training, inequitable implementation, and lack of support or accommodations for low-performing students can all contribute to this widening achievement gap. Demonstrating the impact of technology-based programs will help build public trust in and awareness of effective tools and ensure all students have access to high-quality educational technology that can positively impact their learning and achievement.

Implementing educational technology programs can be expensive and time-consuming and requires specialized training and professional development. The Elementary and Secondary Education Act (ESEA) encourages state and local educational agencies to prioritize evidence-based decisions on using educational technologies in schools. In addition, the Office of Education Technology developed a toolkit to support educational leaders in using evidence to inform the selection of educational technology in schools (Office of Educational Technology, n.d.). Findings from this study contribute to an evidence base that can help schools, districts, and states make informed decisions regarding resource allocation and adoption of technology-based programs, as well as the specific skills and knowledge that educators need to effectively integrate the technology into their schools. This can potentially lead to improving student performance while saving time and money in the long run. By prioritizing evidence-based practices, schools and districts can track the impact of the technologies they adopt and hold themselves accountable for meeting their own goals.

References

Bacher-Hicks, A., Goodman, J., Mulhern, C. (2020). *Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time*. (Working Paper 27555). National Bureau of Economic Research.

Bohrnstedt G., Kitmitto S., Ogut B., Sherman D., Chan D. (2015). *School composition and the Black–White achievement gap* (NCES 2015-018). National Assessment of Educational Progress, National Center for Education Statistics, and U.S. Department of Education.

Center for Research on Education Outcomes. (2020, October). *Estimates of learning loss in the 2019–20 school year*. CREDO at Stanford University. https://credo.stanford.edu/reports/item/report-0/

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). *COVID-19 and learning loss—Disparities grow and students need help*. McKinsey & Company. https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2021). *COVID-19 and education: The lingering effects of unfinished learning*. McKinsey & Company. https://www.mckinsey.com/industries/education/our-insights/covid-19-and-education-the-lingering-effects-of-unfinished-learning

Engzell, P., Frey, A., & Verhagen, M. D. (2020, October 29). *Learning loss due to school closures during the COVID-19 pandemic*. SocArXiv Papers. https://doi.org/10.31235/osf.io/ve4z7

Fry, R., Kennedy, B., & Funk, C. (2021). *STEM jobs see uneven progress in increasing gender, racial and ethnic Diversity*. Pew Research Center. https://www.pewresearch.org/science/wp-content/uploads/sites/16/2021/03/PS_2021.04.01_diversity-in-STEM_REPORT.pdf

Guryan, J., Ludwig, J., Bhatt, M., Cook, P. J., Davis, J. M. V., Dodge, K., Farkas, G., Fryer, R. G., Jr., Mayer, S., Pollack, H., Steinberg, L., & Stoddard, G. (2023). Not too late: Improving academic outcomes among adolescents. *American Economic Review*, *113*(3), 738–765. https://doi.org/10.1257/aer.20210434

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, *24*(4), 470–497. https://doi.org/10.1007/s40593-014-0024-x

Heritage, M., & Popham, W. J. (2013). *Formative assessment in practice: A process of inquiry and action*. Harvard Education Press.

Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science*, *342*(6161), 935–937.

Kraft, M. (2019). *Interpreting effect sizes of education interventions*. (EdWorkingPaper: 19-10). Annenberg Institute at Brown University. http://www.edworkingpapers.com/ai19-10

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549–565. https://doi.org/10.3102/0013189X20965918

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. (NCSER 2013-3000). National Center for Special Education Research and U.S. Department of Education. https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf

Maldonado, J. E., & De Witte, K. (2020). *The effect of school closures on standardised student test outcomes*. (Discussion Paper Series DPS20.17). Katholieke Universiteit (KU) at Leuven.

Morgan, I. (2022*). Equal is not good enough: An analysis of school funding equity access across the U.S. and within each state.* The Education Trust. https://edtrust.org/wp-content/uploads/2014/09/Equal-Is-Not-Good-Enough-December-2022.pdf

Murphy, R., Roschelle, J., Feng, M., & Mason, C. A. (2020). Investigating efficacy, moderators and mediators for an online mathematics homework intervention. *Journal of Research on Educational Effectiveness*, *13*(2), 235–270. https://doi.org/10.1080/19345747.2019.1710885

National Assessment of Educational Progress. (2022). *Largest score declines in NAEP mathematics at grades 4 and 8 since initial assessments in 1990*. National Center for Education Statistics, Institute of Education Sciences, and U.S. Department of Education. https://www.nationsreportcard.gov/highlights/mathematics/2022/

North Carolina Department of Public Instruction. (2022a, April 26). *Historical trends and results 2020–21*. https://www.dpi.nc.gov/historical-trends-and-results-2020-21

North Carolina Department of Public Instruction. (2022b, August 25). *North Carolina end-of-grade tests of mathematics grades 3–8*. https://www.dpi.nc.gov/documents/accountability/testing/eog/eog-mathematics-achievement-level-ranges-and-descriptors

Office for Civil Rights. (2021). *Education in a pandemic: The disparate impacts of COVID-19 on America's students*. U.S. Department of Education. https://www2.ed.gov/about/offices/list/ocr/docs/20210608-impacts-of-covid19.pdf

Office of Educational Technology. (n.d.). *Supporting schools' use of evidence to guide EdTech adoption*. U.S. Department of Education. https://tech.ed.gov/evidence/

Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online mathematics homework increases student achievement. *AERA Open*, *2*(4). https://doi.org/10.1177/2332858416673968

Roschelle, J., & Hodkowski, N. M. (2020). *Using research on homework to improve remote and hybrid learning*. [White Paper]. Digital Promise. http://hdl.handle.net/20.500.12265/107

Sahni, S. D., Polanin, J. R., Zhang, Q., Michaelson, L. E., Caverly, S., Polese, M. L., & Yang, J. (2021, January). *A What Works Clearinghouse rapid evidence review of distance learning programs*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/Distance_Learning_RER_508c.pdf

Schneider, M. (2022). *Downward trends: Pre- and post-pandemic NAEP results*. Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/director/remarks/09-01-2022.asp

Stein, P. (2020, October 30). Washington Post: In D.C., achievement gap widens, early literacy progress declines during pandemic, data show. *The DC Line: News, Views and Events*. https://thedcline.org/2020/10/30/washington-post-in-d-c-achievement-gap-widens-early-literacy-progress-declines-during-pandemic-data-show/

Tipton, E. (2014). How generalizable is your experiment? An index for comparing experimental samples and populations. *Journal of Educational and Behavioral Statistics*, *39*(6), 478–501. https://doi.org/10.3102/1076998614558486

Note about this study: This study was conducted independent of Worcester Polytechnic Institute and The ASSISTments Foundation, who are the founder and distributor of ASSISTments. A local coach was hired by Worcester Polytechnic Institute to provide the professional learning and support to the teachers in the intervention group that was part of the intervention package. All data collection and analysis activities were carried out independently by the study team.